

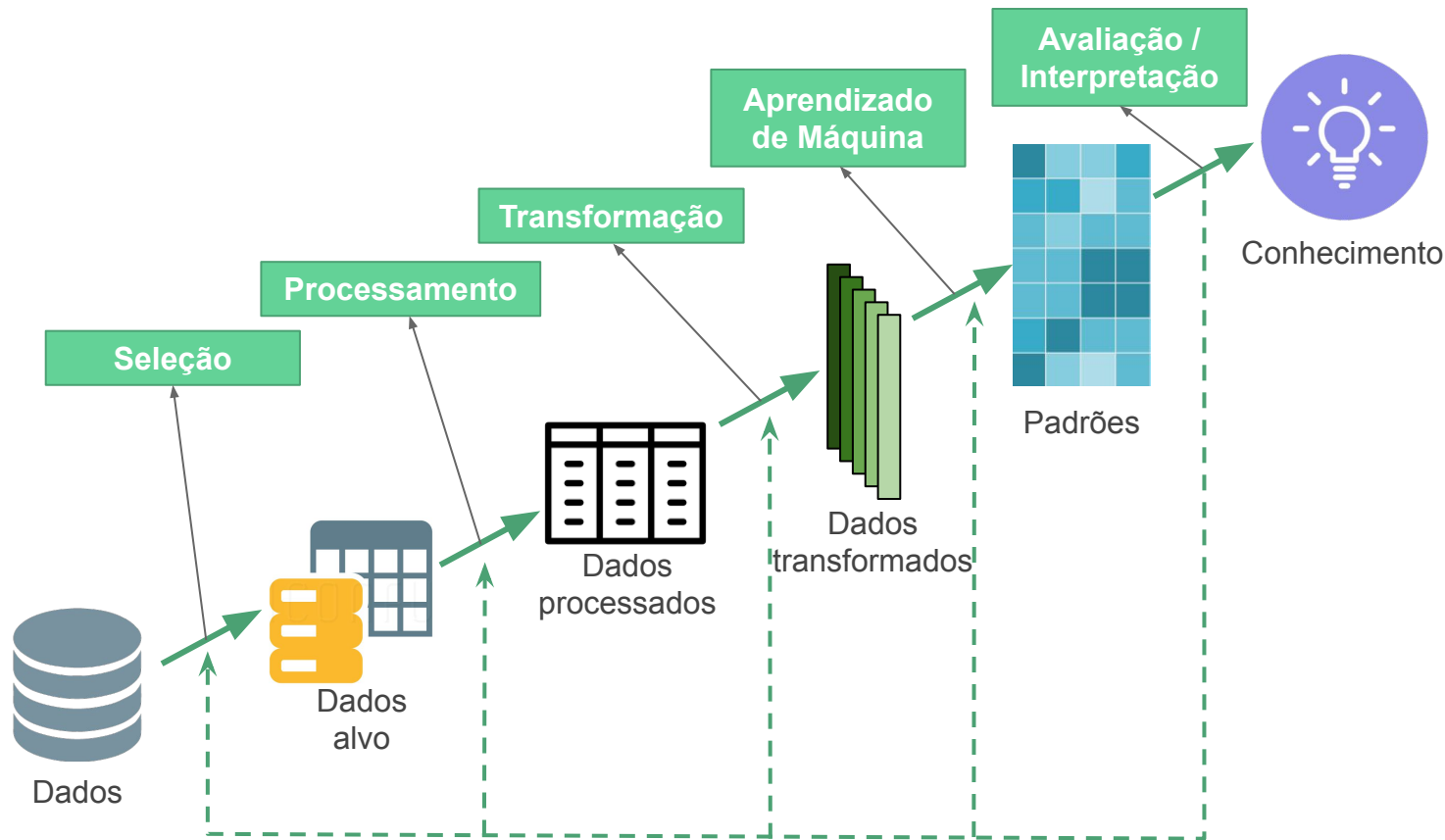
Análise de séries temporais

por similaridade e alinhamento não
linear com Dynamic Time Warping



Tarefas de mineração de séries temporais parte 1

Etapas da descoberta de conhecimento



Do ponto de vista de “data analytics”

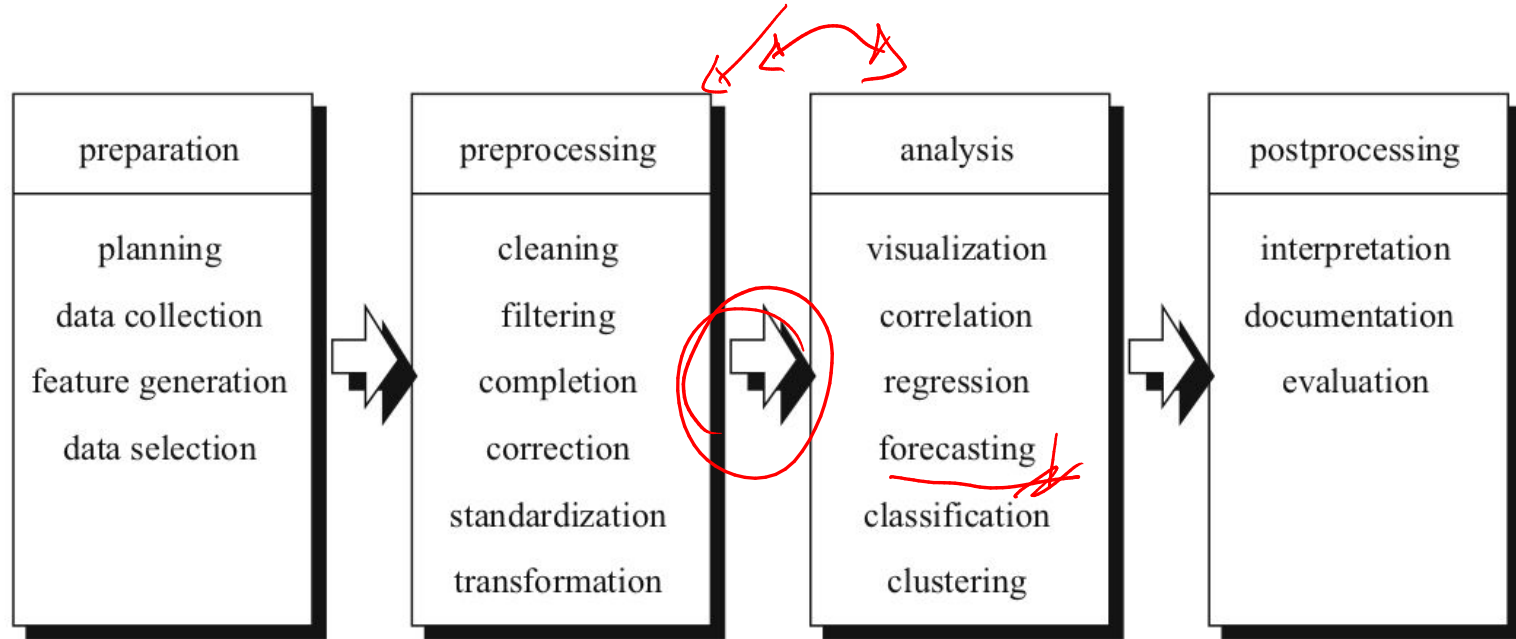


Fig. 1.1 Phases of data analysis projects

Forecasting

Componentes de uma série temporal

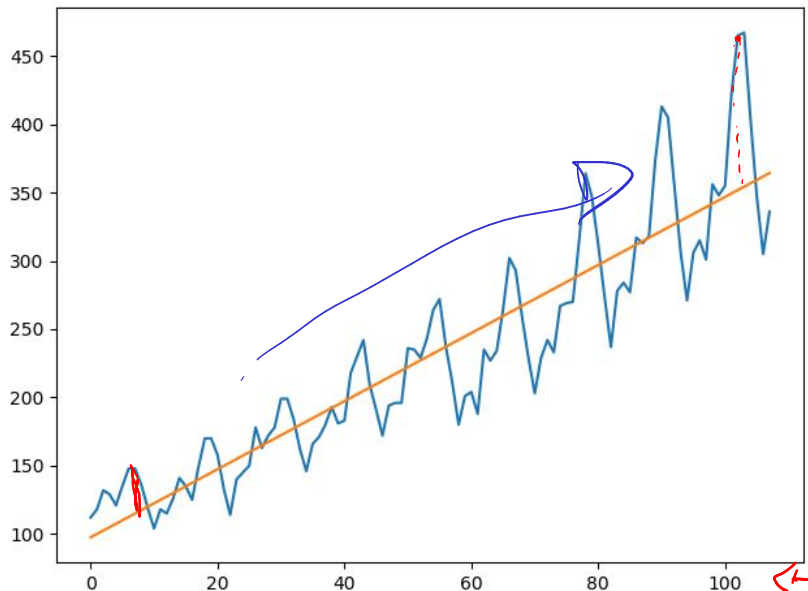
- Tendência ✓
- Sazonalidade ✓
- Ciclicidade ✓
- Ruído (aleatoriedade) ✓

Um conceito importante que é impactado por isso: **estacionariedade**

- Quando uma subsequência $(x_t, x_{t+1}, \dots, x_{t+m})$ tem a mesma distribuição conjunta de $(x_{t+j}, x_{t+j+1}, \dots, x_{t+j+m})$, para qualquer j inteiro
- Estacionariedade de tendência

Componentes de uma série temporal

- **Tendência**
- Sazonalidade
- Ciclicidade
- Ruído (aleatoriedade)

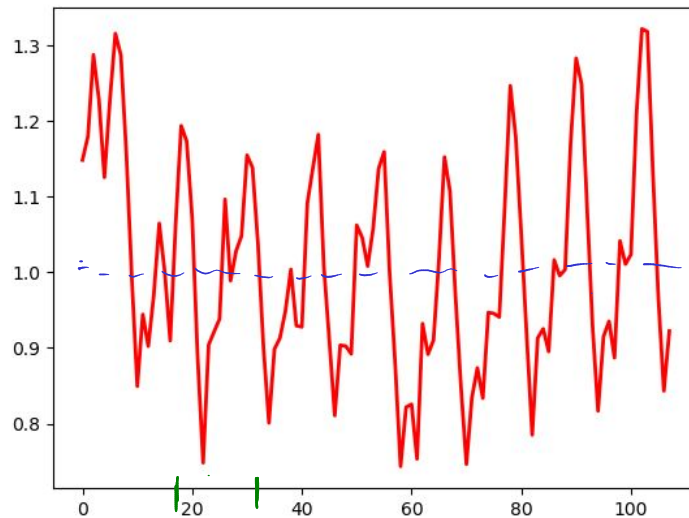


Componentes de uma série temporal

- **Tendência**
- **Sazonalidade**
- Ciclicidade
- Ruído (aleatoriedade)

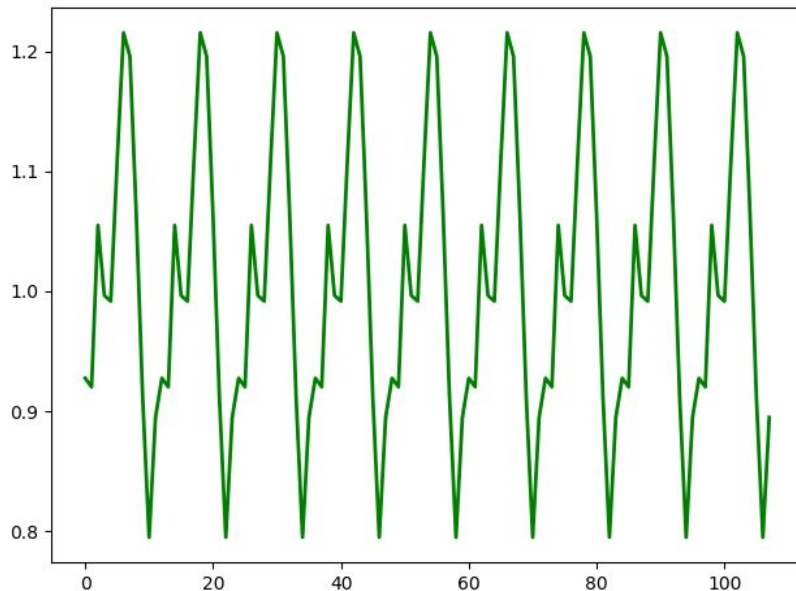
Dickey-Fuller

Dados sem a tendência (X/trend)



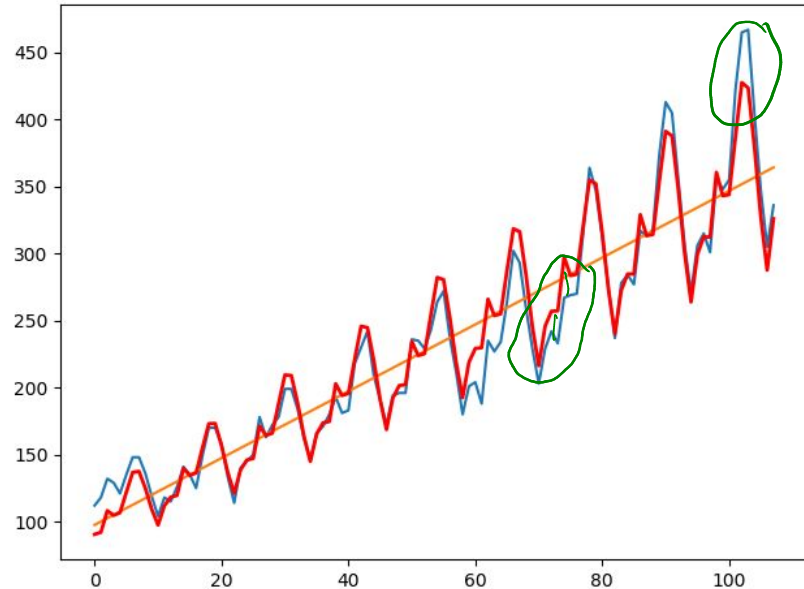
Componentes de uma série temporal

- **Tendência**
- **Sazonalidade**
- Ciclicidade
- Ruído (aleatoriedade)



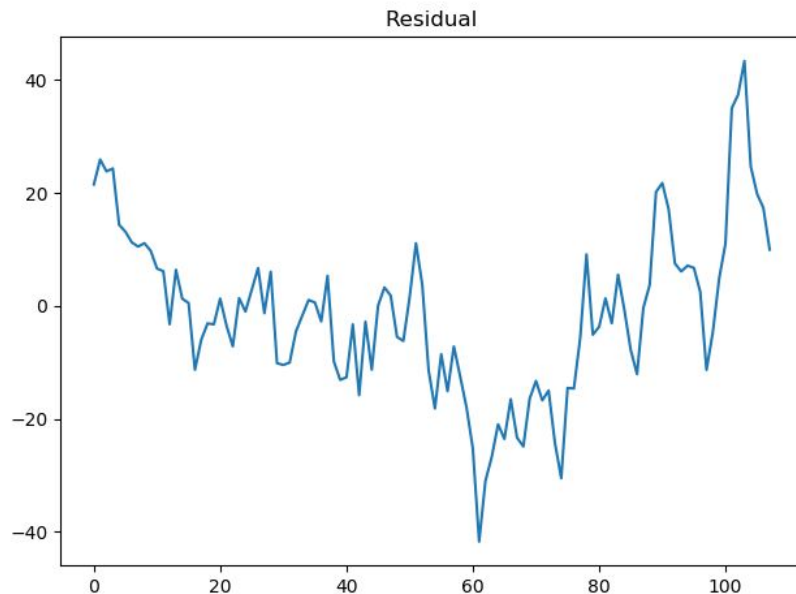
Componentes de uma série temporal

- **Tendência** * (se multiplicativo)
- **Sazonalidade** = aproximação da série
- Ciclicidade
- Ruído (aleatoriedade)



Componentes de uma série temporal

- **Tendência**
- **Sazonalidade**
- Ciclicidade
- Ruído (aleatoriedade)



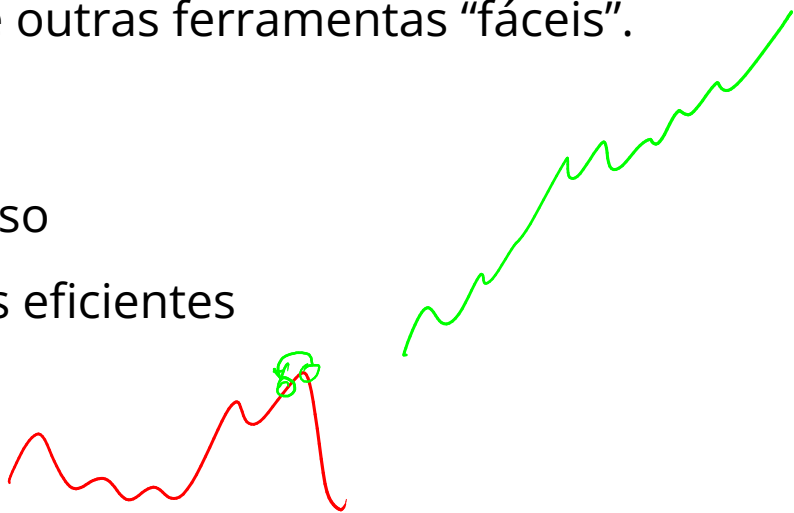
Componentes de uma série temporal

Acabamos de **decomp**or a série temporal

Porém, eu fiz isso usando regressão linear e outras ferramentas “fáceis”.

- Há ferramentas muito melhores para isso
- Há ferramentas até mais simples e mais eficientes

- Ex: pela derivativa -> $y_{t+1} = x_{t+1} - x_t$



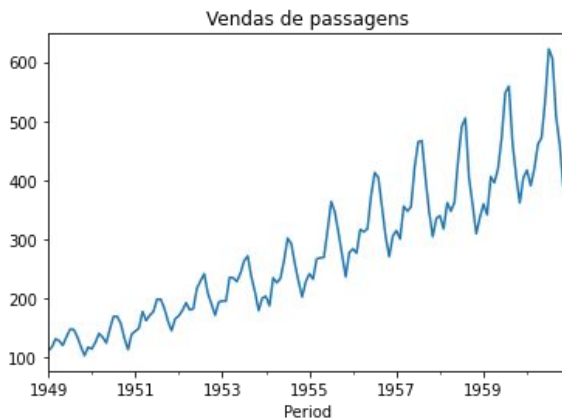
Tirando a tendência

O problema é que muita coisa pode mudar

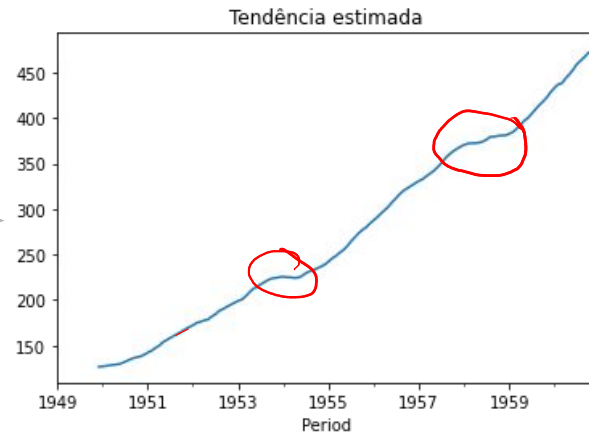
- Tendência global vs. local
- Tendência linear vs não-linear

Mas há coisas simples para trabalhar que são boas o suficiente para a maioria dos casos. Por exemplo, média móvel!

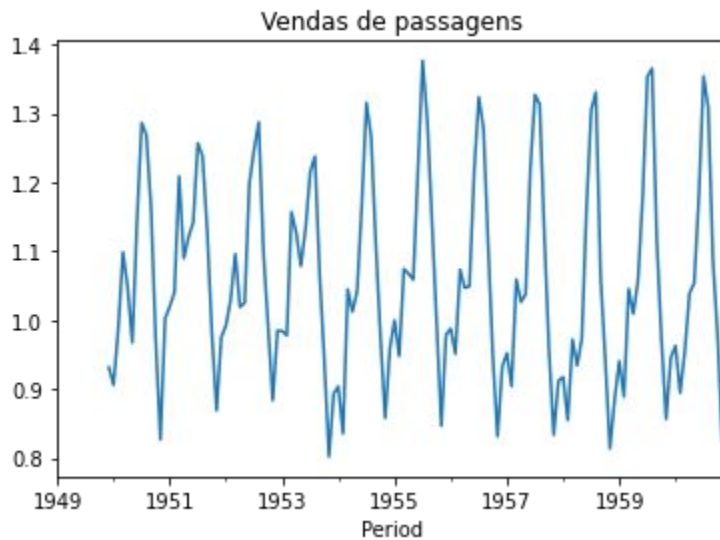
Tirando a tendência



```
y.rolling(12).mean().plot  
(title='Tendência estimada')
```

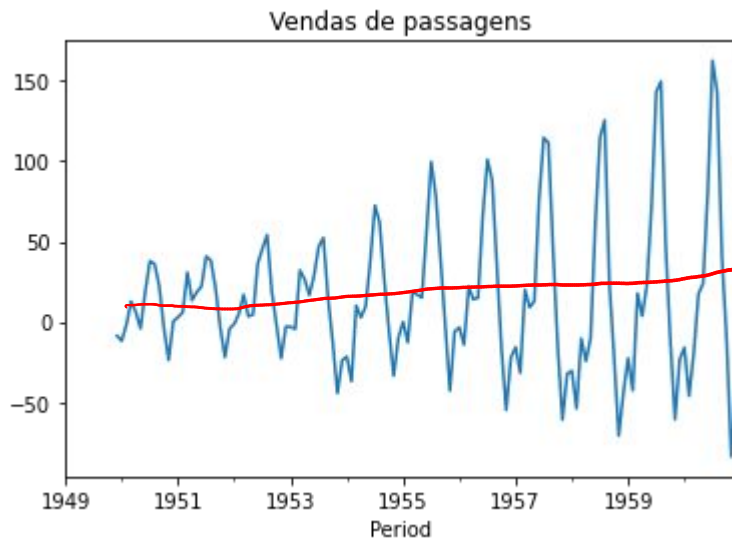


Tirando a tendência



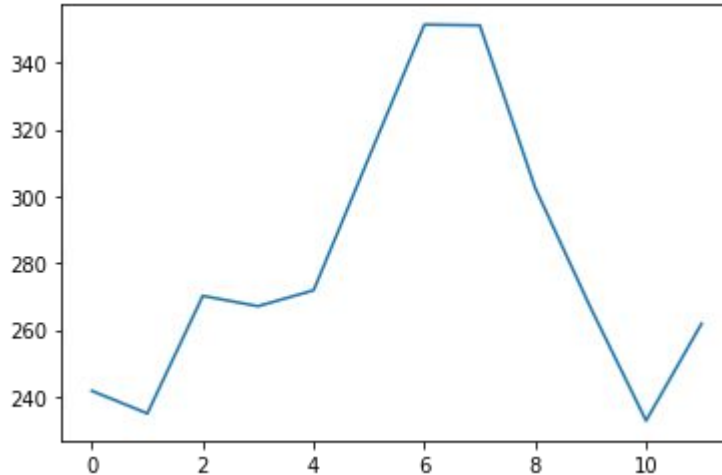
Tirando a tendência

Apenas por curiosidade, se eu tivesse considerado a tendência como aditiva



Verificando a sazonalidade

Ao tirar a média das observações agrupadas a cada 12 observações



Decomposição da série

Para forecasting, a decomposição foi só uma “observação”

- Métodos de forecasting fazem suposições implícitas sobre isso

Mas a decomposição de séries pode ser muito boa para análise em geral

Pare um pouquinho

Forecasting

$$X = (x_1, x_2, \dots, x_n) \leftarrow$$

$$\hat{y} = \hat{x}_{n+t}$$

$t = \text{horizonte} \in \mathbb{N}$

Forecasting - técnicas

Muitas técnicas de forecasting são baseadas em uma soma ponderada das últimas observações.

Vamos começar com o Exponential Smoothing.

Single/Simple Exponential Smoothing (SES)

O SES é apenas uma média das observações passadas, mas com um decaimento exponencial

$$\hat{x}_{t+1} = \sum_{i=0}^t \alpha(1 - \alpha)^i x_{t-i}$$

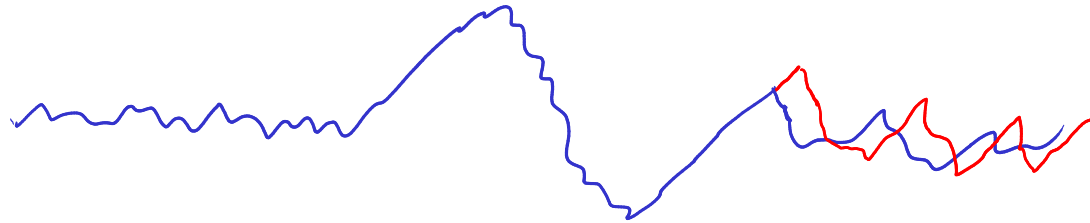
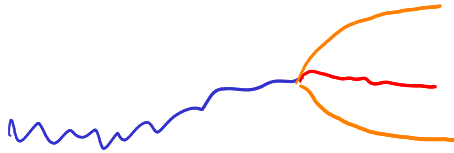
ou seja

$$\alpha \in [0, 1]$$

$$\hat{x}_{t+1} = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + \alpha(1 - \alpha)^3 x_{t-3} + \dots$$






Single/Simple Exponential Smoothing (SES)

- Alpha (α) é um hiper-parâmetro/configuração do usuário
 - Algo perto de 1 dá bastante ênfase para as observações recentes
 - Algo próximo a 0 “distribui” a importância das observações ao longo de toda a série
- Não é adequado para séries com tendência e sazonalidade

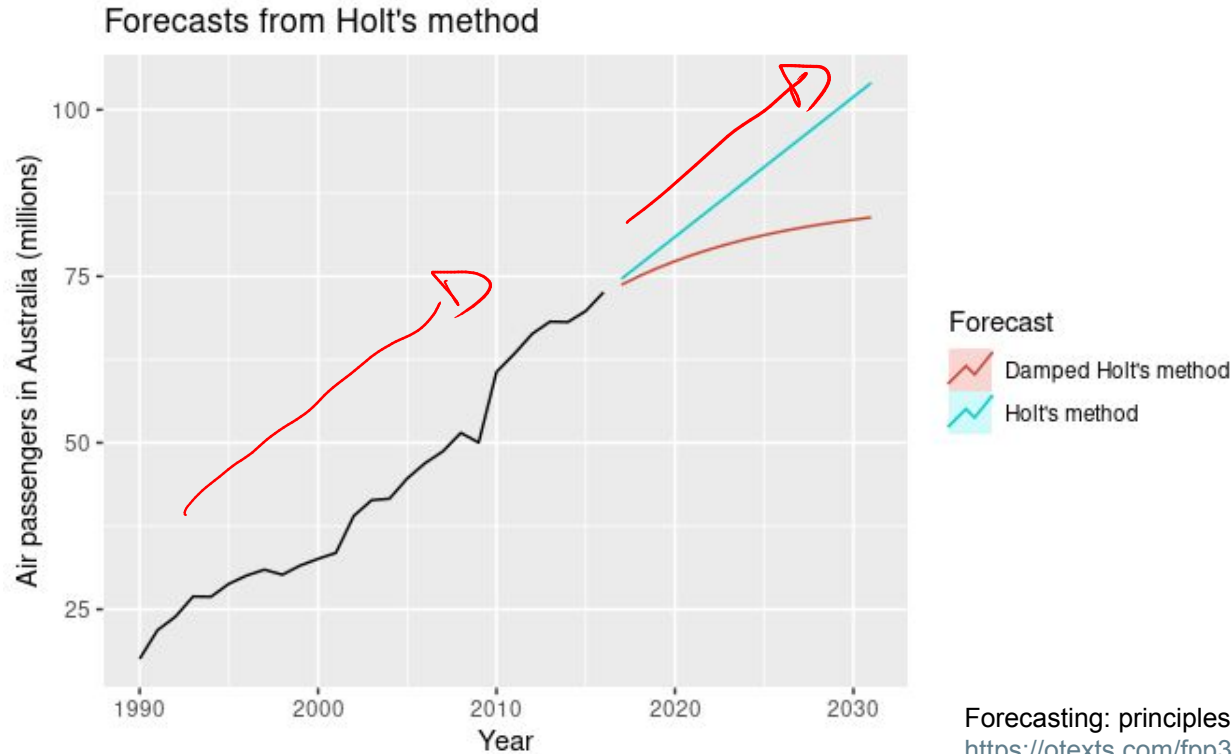


Double Exponential Smoothing (DES)

Adiciona suporte a tendência linear ao SES

- Adiciona um hiper-parâmetro **beta (β)** para controlar o **decaimento** da influência **da tendência** 
- Se considerar tendência linear, chamamos de modelo de tendência linear de **Holt** 
 - Mas, pode considerar tendência multiplicativa
- Para não considerar que a tendência é constante, adiciona um parâmetro de **amortecimento phi (ϕ)**   

Double Exponential Smoothing (DES)



Triple Exponential Smoothing (TES)

Adiciona suporte a sazonalidade

- Adiciona um hiper-parâmetro **gamma (γ)** para controlar o **decaimento** da influência **da sazonalidade**
- **Holt-Winters Exponential Smoothing**
- Além disso, o período deve ser adicionado (parâmetro)

Exponential Smoothing

$> 0,5$

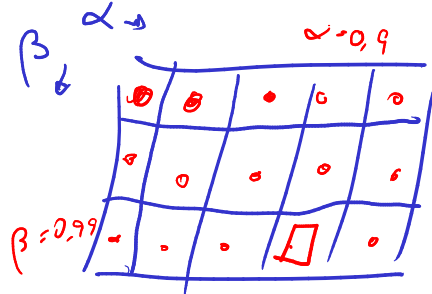
Um dos problemas dessa abordagem é o número de hiper-parâmetros introduzidos

- Podemos usar técnicas de HPO (*hyper-parameter optimization*)

$$\alpha \in (0,01; 0,1; 0,2; \dots 0,99) \quad \subset [0,01; 0,99]$$

$$\beta \in (\dots)$$

Grid search



ARMA

Modelo ARIMA

AutoRegressive Integrated Moving Average

Não está diretamente relacionado a tendência e sazonalidade, mas a autocorrelação nos dados

- Auto Regressão: o valor de uma variável depende somente de observações passadas dela mesma (e algum fator aleatório)
- Integrada: usa diferenciação (ex: subtrair valores consecutivos) para forçar estacionariedade
- Moving Average: usa a dependência entre observação e erro residual da média móvel

SARIMA - ARIMA sazonal

Modelo ARIMA

Auto Regressão

Modelo regressivo de ordem p

- $AR(p)$

$$\begin{aligned} &y_{t-12} \\ &y_{t-24} \\ &y_{t-36} \\ &\vdots \end{aligned}$$

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

onde ε_t é ruído branco e c é o *drift*.

E vez de tentar prever a série em si, tenta-se modelar relações entre observações consecutivas

Modelo ARIMA

Diferenciação

- De primeira ordem

$$y'_t = y_t - y_{t-1}$$

- De segunda ordem

$$\begin{aligned} y''_t &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2}. \end{aligned}$$

- Sazonal

$$y'_t = y_t - y_{t-m}$$

Modelo ARIMA

Modelos *moving-average*

MA(q)

$$y_t = \underline{c} + \underline{\varepsilon}_t + \underline{\theta}_1 \underline{\varepsilon}_{t-1} + \underline{\theta}_2 \underline{\varepsilon}_{t-2} + \cdots + \underline{\theta}_q \underline{\varepsilon}_{t-q},$$

ARIMA

O ARIMA nos deixa com três importantes hiper-parâmetros

- p - ordem do termo AR
- q - ordem do termo MA (tamanho da janela deslizando)
- d - ordem da diferenciação

ARIMA (p, d, q)

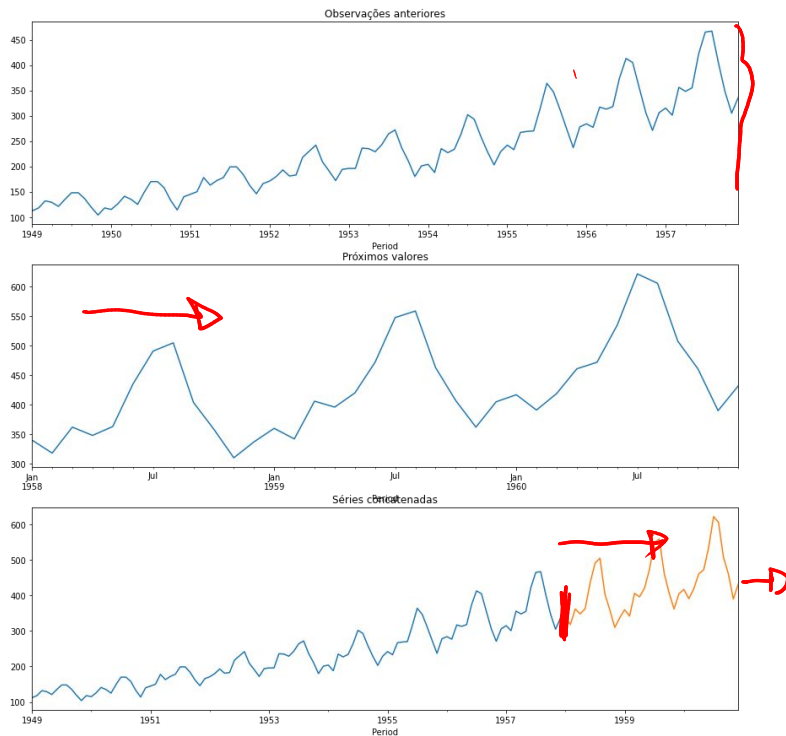
Como testar um algoritmo antes de usar na prática?

É preciso fazer um *holdout* temporal:

- Uma janela de 1 a $n-t-1$ para ajustar seu *forecaster*
- observações de $n-t$ a n para testar

$$MSE = \frac{\sum (\hat{x} - x)^2}{t}$$

Temporal cross-validation



Como testar um algoritmo antes de usar na prática?

- Então vamos testar, na prática
- Chama o Luiz, produção